Culture Counts

BY ANTHONY CIAVARELLI

The importance of establishing and maintaining a positive safety culture and climate in any aviation organization is now beyond debate. But little attention has been paid to measuring an organization's safety environment, an omission that is important because, as business schools preach, you can't manage what you can't measure.

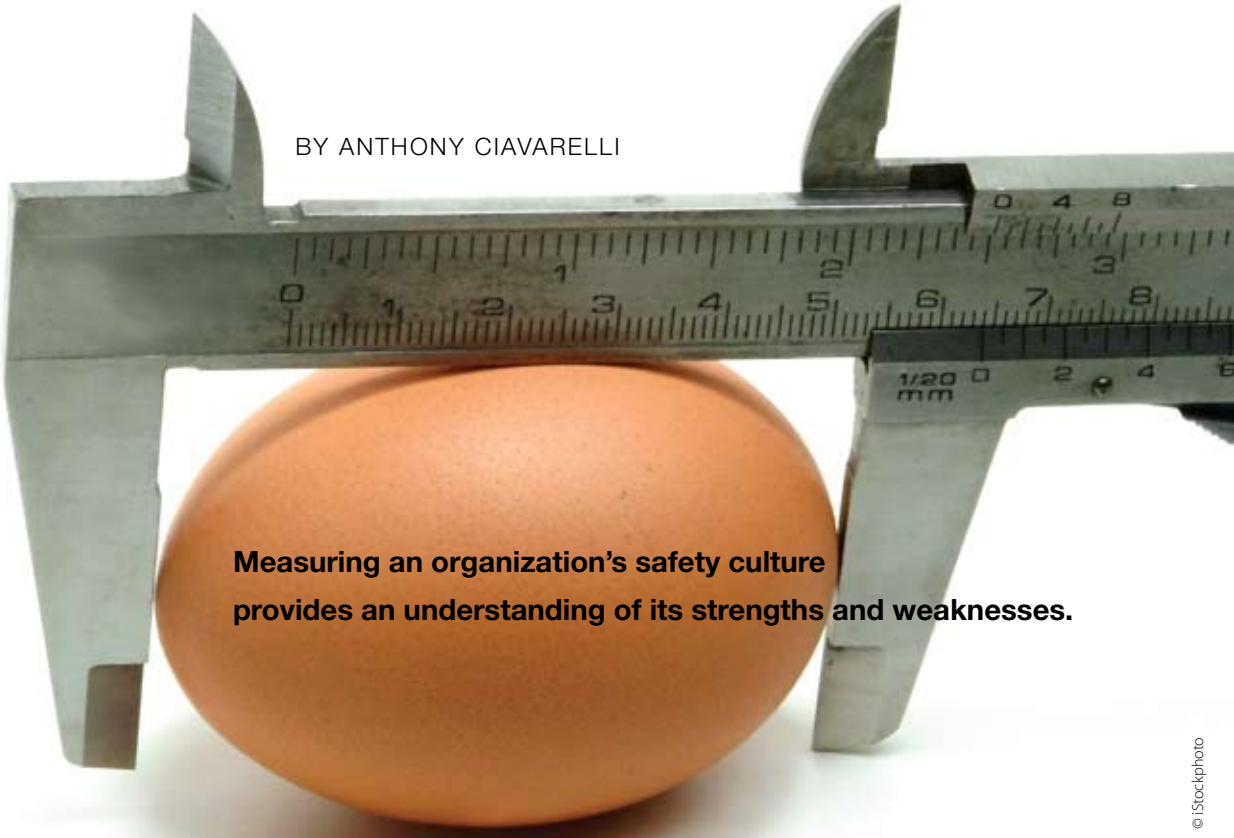
However, an assessment tool developed originally for U.S. Navy aviation units now can provide the foundation for a process of measuring and tracking an organization's safety culture.

Awareness of the existence of safety culture in aviation, and its importance, evolved over recent years through the examination of high-profile catastrophes. The first major airline accident attributed in part to organizational factors was the January 1982 Air Florida Boeing 737 crash in Washington.

The flawed decision chain that led to the accident was a consequence of the company's failure to give the flight crew adequate training, the U.S. National Transportation Safety Board (NTSB) found. This training gap led to pilot judgment error, inappropriate procedures and a breakdown in flight crew communication. This accident was among several that led airlines to develop crew resource management training.

Similarly, the chain of errors that led to the May 1996 ValuJet Airlines McDonnell Douglas MD-80 crash in Florida is one of several examples of an "organizational accident," an accident deeply rooted in a company's lack of leader commitment and support for safety.

The term "safety culture" first appeared in the accident investigation report published by the International Atomic Energy Agency's (AEA) Nuclear Advisory Group following the April 1986 meltdown and steam explosion of the nuclear power plant near Chernobyl, Ukraine. The AEA concluded that the nuclear reactor was poorly designed and the people operating the plant were not properly trained or supervised. The accident, said the World Nuclear Association, was a direct consequence of Cold War isolation and the resulting lack of a safety culture.

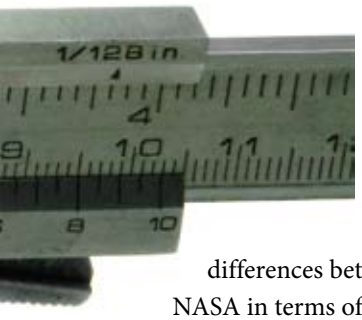


Measuring an organization's safety culture provides an understanding of its strengths and weaknesses.

© iStockphoto

Among the first to bring the term "culture" to the aviation community was John Lauber when he was an NTSB member. Recognizing the influence of organizational factors as the root cause of some aviation accidents, he said of the September 1991 Britt Airways Brasilia accident in Texas, "a probable cause … was the failure of this airline's senior management to establish a corporate culture that encouraged and enforced adherence to approved maintenance and quality assurance procedures."

More recently, the Columbia space shuttle accident investigation linked safety culture with the closely related concept of the "high-reliability organization" (HRO). The Columbia Accident Investigation Board (CAIB) concluded that the shuttle's breakup upon re-entry in February 2003 had as much to do with the organizational culture and structure of the U.S. National Aeronautics and Space Administration (NASA) as it did with the foam chunks that detached from the fuel tank and damaged the shuttle's heat shield during the launch.



The CAIB determined that HRO concepts would be extremely useful in describing the culture that should exist at NASA. The CAIB discussed differences between the Navy and NASA in terms of safety culture and operation as an HRO, and concluded that NASA could substantially benefit by following the Navy's example of best practices.

The most egregious aspect of the organizational accident probably is the failure of management to recognize the signs of an impending disaster. For example, ValuJet's maintenance even before the Florida accident was under scrutiny by the U.S. Federal Aviation Administration. And the Columbia accident was believed to have involved a continuation of the poor safety culture that had been revealed after the 1986 Challenger accident. In both cases, NASA employees in the working ranks had warned their supervisors about the risk of losing a shuttle crew because of known system design flaws.

Safety Culture and Aviation

In spite of the professional's use of the term and many popular notions about it, there is no widely accepted definition of "safety culture." Because there is no common metric for measuring the strength of a particular safety culture, there is no clear method by which an organization can assess its safety culture or diagnose its particular strengths and weaknesses. This results in some frustration on the part of the executives who must manage organizations that necessarily operate in hazardous environments.

The Navy responded to the challenge of measuring and managing safety culture in 1996, when its aviation squadrons experienced a rash of accidents attributed to human factors. One accident, in particular, captured the Navy's attention. An F-14 crashed on takeoff from an airport in Tennessee, killing both crewmembers and several civilians on the ground. Investigators pointed to the failure of commanders to manage the pilot, who was known to take unnecessary risks.

The accident occurred during a cross-country flight, which included a visit to the pilot's home town. With his relatives watching, the pilot attempted a dangerously high-angle takeoff, flew into low clouds, became spatially disoriented and crashed into a residential complex.

Following the F-14 accident, I served as a member of a blue-ribbon panel formed to study the underlying causes of naval aviation aircraft accidents. The panel recognized that, in the Tennessee accident and others like it, investigation boards would find known circumstances that produce risks that were not appropriately managed by the commanders.

The panel's review concluded that accidents of this type are very similar to civilian accidents like ValuJet, Chernobyl, Challenger and Columbia. This finding led to the development of the initial survey instrument designed to assess safety climate, safety culture and related organizational factors.

I also was part of a group from the U.S. Naval Postgraduate School that explored innovative methods for assessing organizational factors, including safety climate and safety culture. We worked with Professor Karlene Roberts of the University of California at Berkeley to construct an employee survey based on her theory of an HRO.

Roberts and her colleagues believe that some organizations operate more reliably than others because they place a higher value on safety and a greater focus on avoiding failure. Roberts conducted field studies on Navy aircraft carriers, at air traffic control facilities and at nuclear power plants, organizations that have learned from experience how to manage their risks. Some characteristics that typify HROs are:

- Accurate perception of hazards and operational risks;
- Commitment and involvement of all management levels in safety;
- Open reporting of unsafe conditions or risk situations;
- Good communication up and down the command chain;

- Continuous training, with high performance standards; and,
- A culture of trust between workers and their supervisors.

The safety climate survey developed on principles of HRO theory is called the Command Safety Assessment Survey (CSA). A Web-based version of the CSA is in regular use by all U.S. Navy and Marine Corps aviation units. Respondents voluntarily and anonymously provide opinions about their organization's safety climate. The similar Maintenance Climate Assessment Survey (MCAS) was developed later for aircraft maintenance personnel. To see an example of the CSA, go to <https://www.hfa-clients.com/flightsafety/login.html>.

Safety Climate and Culture

Later versions of the CSA incorporate various aspects of safety culture and safety climate derived from the work of European social scientists. These researchers greatly improved our understanding of the differences between safety culture and safety climate.

Culture is considered to be the force behind an organization's goals — it drives the means to attain goals and spells out how to achieve success. An organization's cultural values also guide decisions and processes for correcting deviations from norms and expectations.

Safety culture is defined as the shared values, beliefs, assumptions and norms that govern decision making and that may affect individual and group attitudes about danger, safety and the proper conduct of hazardous operations.

Safety climate, an important indicator of the underlying safety culture, refers to the perception of the people in an organization that their leaders are committed to safety, have taken appropriate measures to communicate safety principles and ensure adherence to safety standards and procedures.

The CSA and MCAS surveys, and later applications in civilian aviation, aerospace and health care, are designed to address key aspects of safety climate. Results from the surveys sometimes can be used in the estimation of accident risk based on the extent to which the organization exhibits HRO attributes of leadership commitment to safety, adequacy of policies, adherence to standards and other factors.

The safety climate surveys have been well received in naval aviation and serve as an important source of performance feedback that commanders use to improve the safety of squadron operations.

© Sawayasu Tsuji /iStockphoto

Since its inception in 1997, the CSA has been administered to all Navy and Marine Corps squadrons. There are more than 80,000 survey responses in the Navy's CSA database.

This success led to the development of equivalent online survey systems that are being used worldwide in civilian aviation, aerospace and health care industries.

Over the years, safety climate and culture surveys have matured to provide organizations with reliable and valid measures that produce useful findings. An example (Figure 1) shows the results of the CSA item ratings taken from more than 10,000 respondents, with a comparison across military ranks. The chart shows that much higher ratings are given by senior commissioned officers (lieutenant commanders, commanders and captains) compared with lower-ranking officers (ensigns, lieutenants junior grade and lieutenants). This relationship also holds for the noncommissioned officers (NCOs) — higher ranking NCOs (petty officers) gave higher ratings than the lower ranks.

Ratings by an organization's senior management also have been consistently higher in the civilian aviation, aerospace and health care industries. This suggests that supervisors might not be fully aware of the thoughts and feelings of their subordinates when it comes to the organization's safety climate and the strength of its safety culture.

'Statistical Goodness'

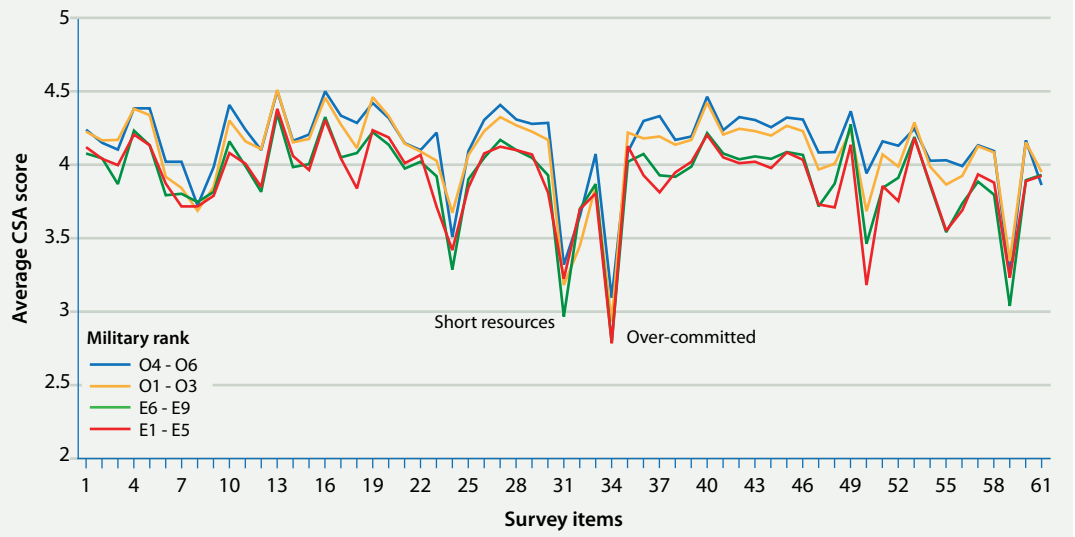
Professional survey developers place a high value on surveys in terms of "statistical goodness." Two of the important aspects of statistical goodness are the survey's reliability and validity.

Basically, when a survey is reliable, the results are relatively error free and consistent — if the survey is administered twice to the same people at the same time, the ratings would be identical.

Unless the survey can be shown to be reliable by using a variety of statistical methods, there is no point in attempting to show that the survey is valid for its purpose. If the ratings are random and unreliable, then the findings cannot be trusted. The statistical reliability of both Navy and civilian surveys was found to be very high.

The term "validity" refers to measuring what we set out to measure — in this case, safety climate and safety culture. The survey instrument must be carefully constructed to reflect the key attributes of climate and culture. This kind of validity is called "content validity" and focuses on the inclusion of

Differences in Average Ratings Across Military Ranks for Survey Items 1-61



CSA = command safety assessment

Source: Anthony Ciavarelli

Figure 1

survey items that reflect some of the underlying organizational dimensions like climate, culture or HRO attributes.

The following are examples of survey statements, selected for adherence to content validity:

- My organization has a realistic view of our operational risks.
- The leadership in this organization is very committed to safe operations.
- All levels of management are actively involved in keeping us safe.
- I am not reluctant to report an unsafe condition or a high-risk incident.
- Deliberate violations of rules or standards are very rare in my organization.
- Sometimes the goal of diagnostic analysis is achieved by organizing survey items into specific measurement areas that reflect different components of climate, culture or HRO attributes. For the CSA, the categories were adapted from studies by UC Berkeley researchers Roberts and Carolyn Libuser.

Five categories, representing different key components of an HRO, are:

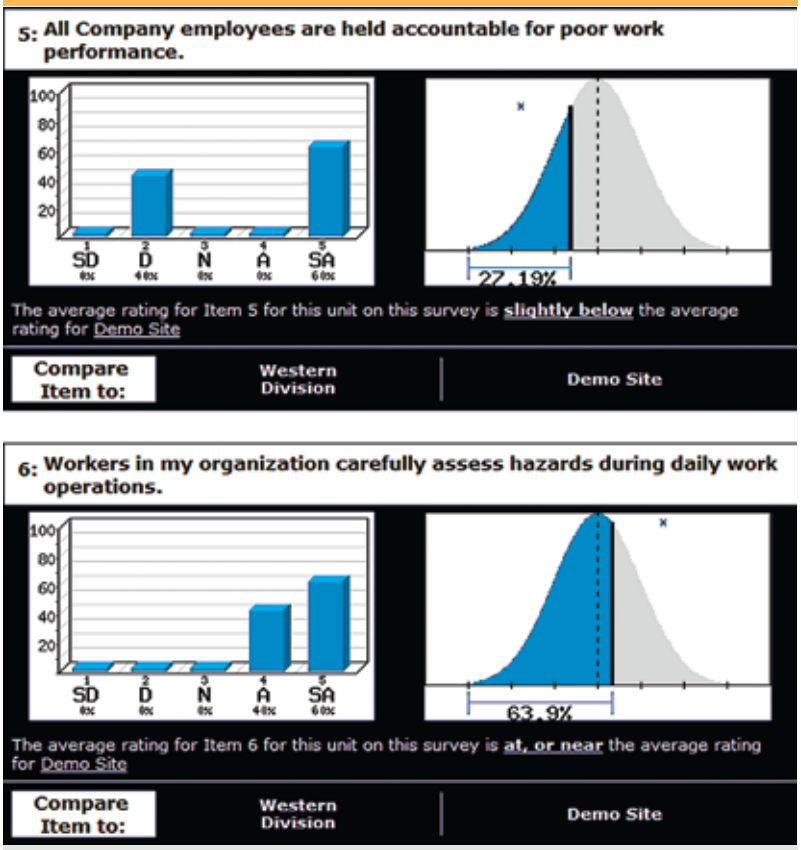
- Safety process auditing — a system of checks and reviews to monitor and improve processes;
- Safety culture and reward system — social recognition that reinforces desired behavior or corrects undesired behavior;
- Quality assurance (QA) — policies and procedures that promote high-quality performance and work performance
- Risk management (RSK MNGT) — whether or not the leaders correctly perceive operational risks and take corrective action; and,
- Leadership and supervision (LDSHP) — policies, procedures and communication processes used to improve people's skills and to proactively manage work activities and operational risk.

Using Survey Data

When completed, a survey's results can be reviewed as an overall outcome, with an average

Supervisors might not be fully aware of the thoughts and feelings of their subordinates when it comes to the organization's safety climate ...

Supervisor's Display Showing Survey Item Results



Source: Anthony Ciavarelli

Figure 2

Normative information can be presented on a supervisor's display (Figure 2). The survey feedback display shows a typical bar chart with agreement percentages along a five-point Likert scale.[1] This display also shows a bell curve indicating the placement of a specific organization's average rating on the bell curve. This placement allows a particular organization — a single department within a company — to compare its average to an overall company average or norm.

Validation Process

Once a reasonable sample of survey data is obtained, another aspect of statistical goodness — predictive validity — can be addressed.

We would expect organizations with a good safety climate and strong safety culture to have a better safety record than those organizations that do not — and this is exactly what we have found in examining the safety climate ratings from surveys taken over the past few years in naval aviation.

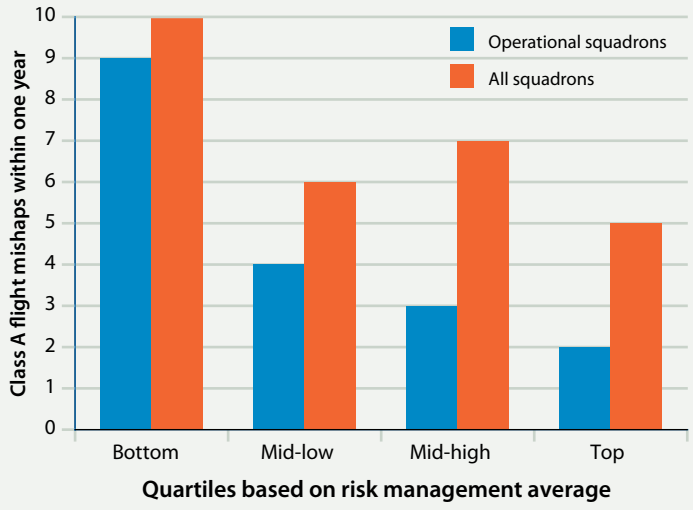
Looking at the relationship between safety climate ratings and safety performance defined in terms of accident frequency we found a much higher number of accidents for low safety climate ratings (Figure 3).

rating across all survey items, or the results can be broken down and compared for each individual HRO component.

The normal distribution curve, or "bell curve" can compare survey results to a specific average or central value in the normal distribution. Using this device, and some statistical computation, an organization can compare its results to an overall average or norm. The norm can be based on a particular company's average rating or on the average for a particular industry.

With this "normative" approach, it would be possible to establish norms for the entire airline industry or another sector of the aviation industry — for example, air traffic control — or beyond aviation to include industry sectors such as aerospace, health care and oil and gas extraction.

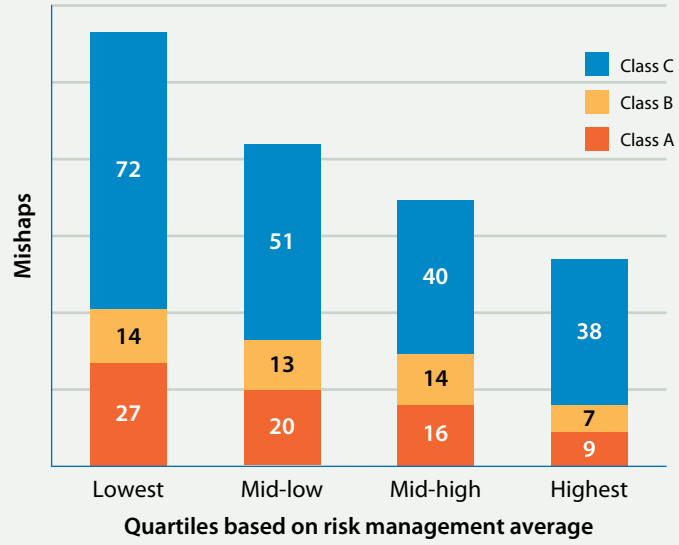
Navy Squadrons' Flight Mishaps by Risk Management Score



Source: Anthony Ciavarelli

Figure 3

Mishap Severity by Risk Management Score

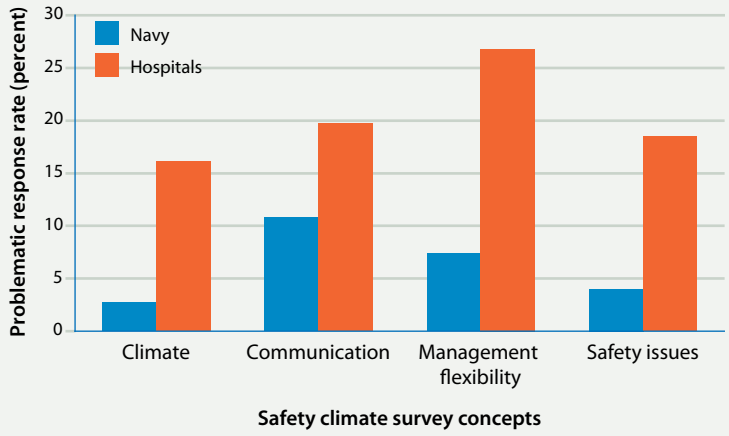


Note: A, B and C are levels of accident severity, "A" being the most severe.

Source: Anthony Ciavarelli

Figure 4

Safety Climate Survey Comparison



Notes:

Source: Anthony Ciavarelli

Figure 5

Accident frequency compared to the CSA risk management subscale also shows a clear relationship between safety climate and safety, with fewer accidents for units scoring higher on safety climate ratings (Figure 4). The units in the lowest quartile of the MCAS had nearly twice the number of accidents (94 versus 49) in the 24-month time frame.

In another application of HRO-based safety climate measurement, we compared ratings obtained from the Navy's CSA to ratings obtained from similar surveys of hospital personnel.

Figure 5 shows an example of the results obtained when 23 common Likert scale survey items used in Navy and hospital studies were compared in terms of the number of "problematic responses," that is, responses that should have been favorable but were not, indicating a fair-to-poor safety climate. The overall problematic response rate was about 6 percent for naval aviation versus 18 percent for hospitals.

A conclusion from the comparison of the Navy and hospital survey responses was that the perception of naval personnel was far more positive because the Navy has had a longer history of focusing on potential failures and has formulated specific processes over the past 60 years or so to ensure that its leadership is active in preventing accidents.

Industries such as aviation, aerospace and health care, as well as the Navy, now recognize the influence of such organizational factors as safety climate and culture on their safety performance. Measuring the state of safety climate and culture, as perceived by employees closest to the daily routines and risk issues, is important to allow managers to keep abreast of hazards and risks inherent in their organization. Survey results provide an organization with the opportunity to identify otherwise unknown risks and to intervene in time to prevent accidents. ●

Anthony Ciavarelli, Ed.D., is a professor at the U.S. Naval Postgraduate School, where he teaches and conducts research in human factors. Ciavarelli founded Human Factors Associates to expand his work to the civilian sector.

Note

1. The Likert scale, developed by psychologist and social scientist Rensis Likert, often is used in surveys. With a five-point scale, respondents are asked whether they strongly disagree, disagree, neither agree nor disagree, agree or strongly agree with survey statements.